CLAIMS

What is claimed is:

1. An apparatus for detecting and handling queue congestion in an output port of a multi-layer network element comprising:

a central processor unit (CPU);

a switching element coupled to the CPU and configured to output packets to a network through the output port, the switching element including:

at least one output queue having storage locations for packet pointers, each pointer configured to point to portions of a packet to be transmitted on the network, associated with the output port, and wherein the number of storage locations is variable;

a start register configured to stored a pointer to the storage location at the front of the queue;

an end register configured to store a pointer to the storage location at the end of the queue as determined by the number of storage location;

a next-free register configured to store a pointer to the next available storage location, wherein packet pointers are stored in the output queue beginning at the location pointed to by the start register and the next-free register is incremented as the next available storage location moves toward the second pointer;

a programmable threshold register configured to store a threshold pointer to a storage location between the location represented by the start register and the location represented by the end register;

threshold logic configured to output a congestion signal when the value in the next free register represents a storage location logically located between the location pointed to key the threshold register and including the storage location pointed to by the end register;

random discarding logic configured to randomly select packets to discard in response to the congestion signal, so that once the threshold is exceed, incoming packets are randomly discarded, using a packet discarding algorithm, such as random early discard, a well known algorithm;

capacity logic configured to output a queue full signal to the CPU when the value in the next free register is equal to the value in the end register;

a memory having at least one entry configured to store information about forwarding decisions for the packet, wherein the entry is adapted to indicate whether packets associated with that entry should be counted;

memory access logic configured to access the entry when an incoming packet associated with that entry arrives at the switching element;

a packet counter configured to count the number of times the entry is accessed, to represent an entry bandwidth;

a computer program mechanism coupled to the CPU configured to compare the contents of the packet counter to a reservation-based protocol negotiated value for lowering a priority of any future packet associated with the entry and destined for the output queue.

2. An apparatus for handling multiple priorities for a multicast packet being output from a network element on at least two output ports comprising:

at least a first output queue and a second output queue, first output queue having a priority higher than the second output queue, at each output port;

a memory configured to output forwarding information about the multicast packet in response to a memory access based in part on a multicast address of the multicast packet, the forwarding information including priority information indicating to which output queue at each output port the multicast packet will be directed.

3. The apparatus of claim 2, further including:

a central processing unit coupled to the memory;

a computer program mechanism coupled to the central processing unit configured to modify the priority information based on an amount of packets being transmitted through one of the output ports.

4. The apparatus of claim 2, further including:

a central processing unit coupled to the memory;

a computer program mechanism coupled to the central processing unit configured to modify the priority information based on information communicated between the network element and an intended reciptent of the multicast packet.

5. An apparatus for queue scheduling in a network element comprising:

at least one output port configured to output packets, each packet having a byte length;

at least two queues associated with each output port, configured to queue packets to be output at each output port;

a weight register associated with each queue and adapted to receive a value representing a weight number;

weighting logic for generating the weight number for each queue;

transmitting logic at each output port configured to transmit packets identified in each queue according to a queue select signal and responsive to a done signal;

scheduling logic at each output port configured to select one of the queues and to generate the queue select signal to the transmitting logic to indicate which queue will be transmitting;

counter logic, at each output port, associated with the counters configured to decrement the weight register equal to a number of bytes transmitted by the transmitting logic;

zero logic configured to transmit the done signal when the number in the counter represents zero;

reloading logic configured to determine the number of packets transmitted after the done signal and to place in the weight register a value equal to the weight number minus the number of packets transmitted after the done signal.

6. The apparatus of claim 5, wherein the scheduling logic is configured to respond to the done signal and the transmitting logic to select a next transmitting queue.

7. The apparatus of claim 5, wherein the scheduling logic is configured to select a next transmitting queue at a time prior to the zero logic generating the done signal.

8. An apparatus in a network element that is adapted to transmit a packet to multiple recipients and includes services for reservation-based protocols, for handling multiple priorities, the apparatus comprising:

at least two output ports, one associated with each of the multiple recipients, each of the output ports having at least a first output queue and at least a second output queue, the first output queue having a priority higher than the second output queue, at each output port;

a memory configured to output forwarding information about the packet in response to a memory access based in part on a header of the packet, the forwarding information including priority information indicating to which output queue at each output port the packet will be directed.

9. The apparatus of claim 8, further including:

a central processing unit coupled to the memory;

a computer program mechanism coupled to the central processing unit and configured to modify the priority information based on an amount of packets being transmitted through one of the output ports.

10. The apparatus of claim 8, further including:

a central processing unit coupled to the memory;

a computer program mechanism coupled to the central processing unit and configured to modify the priority information based on reservation-based protocol information communicated between the network element and an intended reciptent of the multicast packet.